Feasibility of a Semi-computerized Line Bisection Test for Unilateral Visual Neglect Assessment

H. Jee[1], J.Kim[2], C. Kim[3], T. Kim[3], J.Park[2]

[1]Department of Sports and Health Care, Namseoul University, Cheonan-si, Korea [2]School of Information and Communication, Inha University, Incheon-si, Korea [3]Department of Physical & Rehabilitation Medicine, Inha University Hospital, Incheon-si, Korea

Summary

Background: Commonly used paper-and-pencil based test modalities for assessing the degree of unilateral visual neglect (ULN) in patients with hemispheric cerebral lesions consume human resources with a significant inter and intra-rater variability.

Objective: To explore the feasibility of a semi-computerized electronic-pen based ULN assessment system (e-system) to improve assessment quality without altering the conventional user interface.

Materials and Methods: Thirty cognitively healthy participants (HG) and 11 participants diagnosed with right-hemispheric lesion and unilateral visual neglect (NG) were recruited to evaluate the e-system. Line bisection tests (LBT) were repeatedly conducted twice for the inter-rater and intra-rater (reliability) comparisons. The LBT results were assessed by the e-system and the golden standard methods (manual rater assessment). The percent deviation (%), assessment duration (sec), and number of neglected line (each) were evaluated.

Results: The inter-rater comparisons of the assessed deviation (%) variable showed excellent inter-rater reliabilities (CCCs) ranging from .84 (.59 to .95 (p < .001)) to .99 (.90 to .99 (p < .001)) for HG and NG. The Bland Altman mean difference (B-A) plots with bias (95% LOA (limits of agreement)) showed similar agreements between the e-system and the raters ranging from -.04 % (-2.10 to 1.97) to 1.30 % (-2.23 to 4.84) for HG and NG. The effect sizes (ES), which show similarities between the assessment methods, yielded smaller ranges from .01 to .30 for HG and NG. The reliability (test-retest) comparisons showed similar assessment results between the e-system, rater 1, and rater 2. The manual rater assessment time ranging from 5.85 to 6.00 minutes and inter- and intra-assessment variations were virtually eliminated with the e-system.

Conclusion: The semi-computerized system with the conventional paper-and pencil user-interface showed valid and reliable assessment results. It may be a feasible replacement for the manual rater assessment modality even in a clinical setting.

1. Introduction

Advancements in technology have been assisting various fields of medicine as a powerful medium to improve the efficiency and quality of diagnosis and treatment. Assistive technology has a promising future in providing powerful tools for clinicians in the medical field including in the field of cognitive assessment and rehabilitation [1–3]. Assessment and rehabilitation of patients with hemispatial or unilateral visual neglect (ULN) is a specialized fields that requires assistive technology [1]. ULN is characterized by the inability of a person to process and perceive stimuli on one side of the body or environment [4]. ULN is a perplexing neuropsychological syndrome which may be observed in patients with cerebral damage or disease to one side, predominantly the right side (nonspeech dominant), of the brain [5]. It is characterized by an inability to detect, be aware, attend to, or respond to stimuli in spatial locations on the contralateral side of the lesion [5–7]. The patients with ULN show a greater rate of deviation from a given stimuli because their ability to correctly perceive a given stimuli is lost [5,6]. The inability to detect or respond to one side of the body negatively detracts from daily function and quality of life [8,9].

Although the right inferior parietal lobe and superior temporal cortex have been suggested as critical sites for inducing ULN, different anatomical studies showed conflicting results [10]. Therefore, the degree of ULN has been tested through various functional tasks along with conventional clinical tests such as MRI or CT. Traditionally, the assessment of ULN in a clinical setting has involved the use of 'paper-and-pencil' tests [11]. The paper-and-pencil based tests are popular in clinical settings because they are simple and quick to administer. Paper-and-pencil tests such as the line bisection, making copies of a drawing, and target cancellation tests are assessment modalities commonly used to determine the degree of ULN [8,12].

Despite the simplicity of using a paper and a pencil as assessment tools, an extensive amount of literature on the effectiveness and application on ULN exists [8]. However, several drawbacks exist on using the paper-and-pencil based tests. Traditionally, the results of the paper-and-pencil ULN tests have been manually assessed by experienced specialists or assessors in the field of cognitive rehabilitation based on a series of criteria [13]. However, the assessor to assessor individuality and discrepancy may exist to induce intra- and inter-rater variability despite the standardized assessment protocol and years of experience in assessing ULN. In addition, a significant amount of assessor time is spent on the manual assessment of the test results leading to a significant amount of clinical resource consumption [13–15]. Moreover, additional information embedded in the test response to increase the diagnostic accuracy could not be comprehensively acquired through the conventional paper-and-pencil tests [13,14,16]. The conventional manual assessment modality has several limitations. It can only provide a fixed number of data points on a patient's evaluation and allows assessment of the final test results that could be readily identifiable by the naked eyes.

Various computer-based approaches have been suggested to improve the existing ULN assessment and rehabilitation methods [12,13,17]. A few studies using videoed and computer-based test responses have been useful in assessing the dynamic responses of neglect with increased accuracy and resolution [13]. These systems have provided precision in the assessment as well as in the monitoring of a patients' progress [13]. However, there are still some drawbacks with the existing systems. Videoed test responses require additional time and resources for further assessment and timed feedback. Computer-based test systems also may be limited by using different assessment interfaces such as a computer monitors and related tools leading to different cognitive and motor demands [18].

Therefore, a semi-computerized system (e-system) which conciliates the need of the assessors and patients was developed in order to improve the assessment quality without disturbing the user interface. It's a system composed of an electronic pen for writing recognition, a micro-pattern printed paper for the conventional user-friendly interface, and a computer installable real-time recording and assessment program. The e-system was first programmed for the line bisection test (LBT) before programming other test modalities to observe the performance analysis and applicability in patients with ULN [5,6]. The line bisection test (LBT) is a brief and convenient test which allows a generalized detection and quantification of spatial neglect in patients with partial brain damage [5]. The semi-computerized e-system was programmed to assess the same conventional variables such as percent deviation (%), number of neglected lines (each), and test duration (sec) in

order to compare the results with the results of the gold-standard assessment method (manual rater assessment).

2. Objectives

The objectives of this study were to demonstrate the feasibility of a semi-computerized system for ULN assessment (e-system) to improve the assessment quality and assist the assessors without altering the user friendly environment for the ULN patients.

3. Materials and Methods

3.1 Participants

The assessments were conducted with three different groups: a cognitively healthy group (HG), stroke patients diagnosed with a right side brain lesion and unilateral visual neglect (NG), and another cognitively healthy group after recalibration of the e-system (HGm). The study protocol and procedures were approved and performed according to the guidelines set by the institutional medical ethics committee of the Inha university hospital. The committee follows the ethical standards of the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects. Informed consent was obtained in writing from each participant after a detailed explanation of the meaning and procedure of the study.

First, the line bisection test (LBT) was conducted on a group of young individuals for the overall e-system evaluation before applying the system to a group of stroke patients. Forty-two cognitively healthy volunteers (21M; 22F) with a mean (SD) age of 24.65 (2.87) years and an education duration of 16.32 (1.93) years (HG) participated in the study.

After observing the stability of the system, it was applied to a group of cognitively impaired patients with unilateral visual neglect (NG) to observe the applicability in a clinical setting. The participants of the second group (NG) were recruited from a hospital-based neurological rehabilitation center located in Incheon, Korea. Eleven patients (10M; 1F) with a mean (SD) age of 63.54 (14.72) years and an education duration of 11.23 (5.1) years (NG) participated in the study. The patients who underwent extensive diagnosis including MRIs and CTs to verify the nature and extent of their neurologic deficits were referred by the attending physicians who had several years of experience. Inclusion criteria included the following: 1) participation in a rehabilitation program (to ensure sufficient muscular strength), 2) able to follow verbal directions, 3) adjusted to medication, 4) independent in self-care activities, and 5) in a stable clinical and metabolic state at the time of testing. Patients were excluded if the diagnosis was inconclusive or if general health complications precluded testing. In addition, all recruited patients were right-handed and had right hemisphere damage as a result of a stroke. The first diagnosis was made within 6 months before the start of the study.

Another set of LBTs was conducted in a new group of cognitively healthy patients (HGm). Some of the bisecting lines during the bisecting performances were not detected by the e-system. Therefore, the system was modified to increase the sensitivity before applying it to another set of healthy individuals. The participants of the third group (HGm) were recruited from Inha university. Eleven cognitively healthy volunteers (5M; 5F) with a mean (SD) age of 23.6 (2.18) years and an education duration of 16.31 (1.97) years participated in the study.

3.2 Instrumentation: Semi-computerized E-system for ULN assessment

The semi-computerized e-system is composed of a position pattern recognizing electronic pen (e-pen), a micro-pattern printed paper, and an installable line bisection test and assessment program on an ordinary computer (▶ Figure 1). The e-pen, structured as with a ball-point pen detects and sends written or drawn information from a position-coded micro-pattern printed paper to a computer installed with the LBT program. A clinical specialist can view the bisecting lines on a computer screen simultaneously as the participants draws the lines on a given LBT paper (▶ Figure 1A). A

clinical specialist or assessor can save the calculated or assessed LBT results on a spreadsheet (i.e., Microsoft Excel format) and as an image (i.e. JPG) (▶Figure 1E).

▶Figure 1B shows the internal structure of the e-pen which is composed of a pressure sensor, Infra-red LED (light emitting diode), image sensor, DSP (digital signal processor), and Bluetooth interface. When the ballpoint pen tip touches the surface of the paper, the pressure sensor measures the pressure of the pen-touch and enables DSP to start the position recognition algorithm. The infra-red LED lighting captures clear images and eliminates the influences from the ambient lighting and hand position. Various real-time image processing algorithms were adopted to extract the position information. The position information is repeatedly extracted from the captured images 85 times per second and the intermediate trajectory between the recognized positions is calculated using a polynomial regression algorithm. The e-pen wirelessly sends the extracted information over a Bluetooth communication interface instantaneously to a computer for simultaneous recording and assessment.

The precise position of the e-pen is recognized by reading the micro-patterns printed on an ordinary paper. The micro-patterns are composed of several tiny dots (diameter 50~80 μm) unique to specific location of the paper. A group of 16 dots can encode one x-y position. The resolution of each location or position is .1 mm. Such detailed resolution naturally reflects and records any hand writing or drawing information on a computer. Since the location information of one x-y position represents the absolution coordinate, instead of the relative coordination, the absolute position of the e-pen can be recognized regardless of the paper rotation. ▶Figures 1C and 1D show the original position pattern image captured by the image sensor and processed image by DSP, respectively. The system was developed by the Embedded Computing Laboratory of Inha university located in Incheon, Korea.

Twenty lines and dotted micro-patterns were pre-printed with black ink on recommended-sized papers (21.5 × 28 cm) prior to the LBT test. Twenty lines were drawn with black ink on a sheet of white paper parallel to its long axis. 100, 120, 140, 160, 180, and 200 mm lines in length were randomly selected and drawn over 18 lines in between two 150-mm centered lines drawn as references on the top and bottom of the paper (▶Figure 1E) [5]. The dotted micro-patterns are sensed by a sensor located inside the e-pen as the bisecting lines are drawn on a LBT paper for the bisecting information to be recorded and analyzed in real-time. The e-system with the bisecting line location information reads the positional information sent from the e-pen to determine the degree of deviation from the center of each line. The e-pen recognizes the micro-patterns unique to specific location beneath the bisecting lines. Since the dotted micro-patterns are virtually undetectable by naked eyes, the participants can only recognize the printed bisecting lines as they conduct LBT with an e-pen (▶Figure 1).

3.3 Experimental procedures

The line bisection test (LBT) was first conducted on the cognitively healthy participants (HG). The test was administered in an isolated room with a table and two chairs for the examiner and the participant to sit on the opposite side to face each other at the table. The K-MMSE (Korean version of the Mini-Mental Status Examination) was first conducted to assess the cognitive status of the participants prior to LBT. The K-MMSE test duration and scores were recorded. The e-pen and micro-patterned paper with printed bisecting lines were given to the participant and the examiner explained the step-by-step procedure of the test.

The LBT instructions were as follows: 1) use the right hand and keep the other hand of the table 2) cut each line in half by placing a small pen notch through each line as close to its center as possible; 3) do not make more than one mark on any line; 4) mark each of the lines without skipping any; and 5) do not move the test page. The paper was taped to the table directly in front of the participant to avoid excessive movement when necessary. All participants used their dominant right hand while responding [5].

The LB test results were scored in two ways. First, the LBT results sent to the computer were instantaneously calculated to display the results on the computer monitor. Second, the hand-drawn results were given to two blinded raters and scored on two separate occasions. The raters were also clinical specialists in the field of neurological rehabilitation with more than two years of experience

with the patients and ULN assessments. The raters counted the number and position of neglected or unmarked lines and measured the deviation of the attempted bisection of each of the lines from the true center using a metric ruler (cm). The LBT duration (sec), which is the total time for the participants to complete bisecting 18 lines in halves, was automatically recorded by the e-system (▶Table 1). The percent deviation (%), assessment duration (sec), and neglected lines (each) were assessed and recorded separately by the e-system and by two raters (rater 1 and rater 2) for comparisons (▶Table 1).

The percent deviation (%) indicates the average deviation of the attempted bisection from each line's true center. The length of the left or right side of the line was measured to the nearest half millimeter and converted to a standard score. The 18 measured scores were converted to absolute values to obtain a mean percent deviation value for each participant. The final value indicated the degree of deviation without the directional information. Neglected lines indicate the number of lines missed during the test by the participants. Each assessment involved one rater with a metric ruler (cm), pen or pencil, calculator, and an assistant with a timer. The following formula was used to calculate the percent deviation (%) [5]:

Percent deviation (%) = (measured half - true half)/true half *100

All bisected lines which did not cross the printed lines or which did not follow the order of the bisecting lines (i.e, 18 bisecting lines from top to bottom) were all checked as neglected lines. The assessment duration (sec) indicates the time the e-system and each rater took to complete assessing the LBT completed by the participants.

The tests were conducted in a test-retest manner with a minimum 48-hour time lapse between each test to avoid the learning effect for the reliability assessment. During the retest, the LBT test papers with 18 randomly generated lines were blindly given to the participants on their second visit to further avoid the learning effect. Therefore, two sets of LBT were administered on each participant. Each LBT test was assessed by the e-system, rater 1, and rater 2, twice. After conducting the LBT on the cognitively healthy participants (HG), another set of LBTs was administered on the cognitively impaired patients with unilateral visual neglect (NG). Finally, one more set of LBTs was given to another group of cognitively healthy participants (HGm) after the modification to increase the sensitivity of the e-system.

3.4 Statistical analyses

The sample size was first considered from previous studies (n = 10 to 43) regarding the reliability of the line bisection test [6,18,19]. Prior to the analytical assessments of the data sets, the normality analyses were initially conducted using the Komogorov-Smirnov and Shapiro-Wilks tests for HG and NG, respectively (p > .05). Both data sets were normally distributed with a very strong straight line of fit.

One way ANOVA (analysis of variance) was conducted between the e-system, rater 1, and rater 2 on all tests to observe the significant differences between the assessment results (mean ± standard deviation (SD)). A repeated measure ANOVA was applied to the repeated measures (test-retest). The inter-rater (e-system vs golden standard) and intra-rater (test-retest) analyses were conducted on the assessed results (evaluation time, neglected lines, and deviation (%)) by the e-system, rater 1, and rater 2. The inter-rater assessment of the e-system was conducted by comparing the assessment results of the e-system to the manual assessment results of the two raters (golden standard). The intra-rater assessment was conducted by comparing the two separate LBT tests (test-retest) repeatedly conducted by the participants on two separate occasions. Concordance Correlation Coefficients (CCCs) with confidence intervals (CIs) were used to measure the degree of fit or inter-rater reliability of the e-system. Intra-class correlations (ICCs) with confidence intervals (CIs) were used to assess the intra-rater or test-retest reliability. CCCs or ICCs between .60 and .80 were considered 'substantial'; coefficients greater than .80 were regarded as 'excellent' (near perfect) [20].

In addition, a Bland-Altman (B-A) mean difference plot with mean differences (bias) and percent limits of agreement (mean difference ± 1.96 SD standard deviation)) (% LOA) were used to provide a clinical estimation of agreement between all paired comparisons [21]. The B-A mean difference plots indicate the degree of agreement between two clinical measurement methods or compare one

method to a golden standard. If the bias and CI are closer to zero, the rater agreements are considered strong or similar to one another [21].

The effect size (Cohen's D) which quantifies the difference between measurements was also calculated to further strengthen the degree of agreement between the results [22]. An effect size of .2 to .3 is considered to have a 'small' effect, around .5 a 'medium' effect and .8 to infinity, a 'large' effect [23]. The data were collected in a pre-designed case data sheet and analyzed accordingly with Med-CalcR statistical software, Version 12.0 (Mariakerke, Belgium). A p-value of <.05 was used to determine the statistical significance.

4. Results

The feasibility assessment of the e-system was first conducted on a group of healthy participants to observe the intra-rater and inter-rater reliability. After observing the applicability in healthy individuals, the e-system was used to assess cognitively impaired patients with unilateral visual neglect. The mean (SD) K-MMSE score and LBT duration of the healthy group (HG) were 29.2 (± .9) out of 30 points and 213.0 (± 23.5) seconds. The mean K-MMSE score and LBT duration of the patients (NG) were 24.82 (± 3.95) out of 30 points and 429.3 (± 86.5) seconds, respectively. ▶Table 1 presents baseline characteristics of the participants.

4.1 Inter-rater reliability of the novel E-system

The inter-rater assessment of the e-system was conducted by comparing the major assessment variable, the percent deviation (%), between the assessment methods (e-system and gold-standard). The novel e-system was first applied to HG for overall usability before being applied to NG. The analysis of variance, effect sizes (ES), concordance correlation coefficients (CCCs), and B-A mean differences of the percent deviation (%) variable results were assessed by the e-system, rater 1, and rater 2. The e-system, rater 1, and rater 2 did not show significant differences between the mean percent deviation values (p value > .5). Inter-rater comparisons between the e-system and both raters yielded excellent (nearly perfect) observer agreements or CCCs for all test comparisons (e-system vs rater 1, e-system vs rater 2, and rater 1 vs rater 2), varying from .84 (.59 to .95 (p < .001)) to .89 (.66 to .97 (p < .001)) for the first test and .97 (.90 to .99 (p < .001)) to .98 (.92 to .99 (p < .001)) for the two test results (▶Table 1). The effect sizes (ES), which indicate the closeness between the two methods, also showed a small ES or strong agreements between all assessment methods, varying from .01 to .12. Finally, the B-A mean differences (bias (95% LOA)) also showed strong agreements between the comparisons, varying from -.09 % (-1.75 to 1.57) to .24 % (-1.82 to 2.30) (▶Figure 2 (a)-(f)).

After observing the stability of the inter-rater comparison results of the e-system on the healthy participants (HG), the e-system was applied to the cognitively impaired patients with diagnosed unilateral visual neglect (NG) to observe the clinical applicability. Identical assessments were conducted for NG. Analysis of variances showed significant differences between the e-system and rater 2 (5.94 ± 5.28 % vs. 4.49 ± 4.24 % (p = .02)) compared during the first test and the e-system and rater 1 (5.61 ± 5.59 % vs. 4.68 ± 5.51 % (p = .01)) compared during the second test. However, the inter-rater comparisons between the e-system and both raters yield excellent (nearly perfect) observer agreements or CCCs for all tests, varying from 84 (.59 to .95 (p < .001)) to .98 (.92 to .99 (p < .001)). Effect sizes also showed small ES or strong agreements between all assessment methods, varying from .02 to .30 (▶Table 2). The B-A mean differences also showed strong agreements between the comparisons, varying from -.32 % (-2.46 to 1.81) to .40 % (-2.23 to 4.84) (▶Figure 3 (a)-(f)).

4.2 Intra-rater (test-retest) reliability and descriptive statistics

The intra-rater reliability assessment of the e-system was conducted by administering the LBT twice with a 48-hour washout period in between the tests. The results were assessed by the e-system and two raters for the HG and NG groups. Analysis of variance and the intra-rater reliability comparisons of the deviation (%) variable were conducted for the e-system, rater 1, and rater 2 as in ▶Table

1. Tests for HG were conducted first to observe for usability before proceeding with NG. The analysis of variance for repeated measures showed significance between the first and second test for the e-system (2.51 ± 1.86 vs. 3.31 ± 2.58 (p < .001)), rater 1 (2.27 ± 1.95 vs. 3.35 ± 2.54 (p = .001)), and rater 2 (2.36 ± 1.98 vs. 3.20 ± 2.64 (p < .001)). The pair-wise comparisons between the test-retest yielded similar substantial intra-rater reliabilities for the e-system (ICC = .66 (.37 to .82 (p < .001)), rater 1 (ICC = .64 (.34 to .81 (p < .001)), and rater 2 (ICC = .73 (.51 to .86 (p < .001)) (▶ Table 1). The effect sizes (ES), which indicate the closeness between the two methods, also showed small ES or strong agreements between all assessment methods, varying from .03 to .11 (▶ Table 1). Finally, the B-A mean differences (Bias (95% LOA)) also showed strong agreements between the comparisons: e-system (-.73 (-5.31 to 3.86)), rater 1 (-1.00 % (-5.57 to 3.56)), and rater 2 (-.76 % (-4.98 to 3.46)) (▶ Figure 2 (g)-(i)).

After observing the stability of the intra-rater reliability results of the e-system on healthy participants (HG), the e-system was applied to the cognitively impaired patients with diagnosed unilateral visual neglect (NG) to observe the clinical applicability. Identical assessments were conducted for NG. Analysis of variances showed no significant differences between the two repeated tests for the e-system, rater 1, and rater 2. The pair-wise comparisons between the repeated tests showed excellent (nearly perfect) observer agreements or ICCs for all tests, varying from (.84 (.40 to .96), (p < .001)) to .91 (.68 to .98), (p < .001)). Effect sizes also showed small ES or strong agreements between all assessment methods, varying from .02 to .30 (▶ Table 2). The B-A mean differences also showed strong agreements between the comparisons varying from -.52 (-5.85 to 4.81) to -.72 (-6.17 to 4.72) (▶ Figure3 (g)-(i)).

4.3 Comparative results of the assessment duration and neglected line

Rater 1 and rater 2 spent 325.93 (± 161.70) and 360.30 (± 157.87) seconds, respectively, as the assessment duration for the first LBT results, and 315.74 (± 130.13) and 349.36 (± 135.47) seconds, respectively, as the assessment duration for the second tests. The assessment duration for the e-system was virtually zero.

In terms of neglected lines, the e-system assessed mean (SD) of .81 (± 1.05) lines for the first test and .93 (± 1.26) lines for the second test for HG. The e-system assessed mean (SD) of 2.27 (± 2.65) lines for the first test and 2.91 (± 2.55) lines for the second test for NG. Conversely, both rater 1 and rater 2 assessed no neglected lines. The e-system assessed mean (SD) of 2.27 (± 2.65) lines for the first test and 2.91 (± 2.55) lines for the second test for NG. The rater 1 and rater 2 assessed .73 (± 1.68), .73 (± 1.68), 1.00 (± 2.49), and .82 (± 1.94) neglected lines for the first and second tests for NG, respectively. The e-system neglected 2 fold more than both rater 1 and rater 2 for all tests. The e-system showed an average of 2 extra neglected lines per participant.

4.4 Inter-rater assessment for the calibrated e-system

The e-system was recalibrated to increase the sensitivity of the system to eliminate excessive neglected lines. Eleven cognitively healthy participants were recruited again to conduct the inter-rater comparisons between the e-system, rater 1, and rater 2. The number of neglected lines assessed by the e-system, rater 1, and rater 2 were zero. The recalibrated e-system maintained a high level of inter-rater reliability as with the previous results.

5. Discussion

Unilateral visual neglect (ULN) is a neuropsychological condition in which a deficit in attention to and awareness of one side of space is observed after sustained damage to one side of the brain [4]. The ability to correctly recognize a given stimuli is lost in patients with ULN, resulting in greater deviation and neglect from a given point [5,6]. A paper-and-pencil test modality has been widely used to assess and rehabilitate cognitively impaired patients with ULN. The line bisection test (LBT) is one of the testing modalities most commonly used to evaluate and rehabilitate patients with ULN

[11,24]. LBT was programmed into a semi-computerized e-system to observe possible application of the system in a clinical setting to evaluate ULN in this study.

The ULN assessments could be performed using the traditional paper-and-pencil tools or more advance computerized devices. One of the major differences between the two testing modalities is that the paper-and-pencil method uses a sheet of paper to present the bisecting lines and a writing tool such as pencil or pen to bisect the lines, and the computer method uses a monitor to present the lines and a mouse-controlled cursor to bisect the lines [18,25].

Although the traditional paper-and-pencil and computerized modalities are considered to be both manual tasks with identical testing formats and procedures, the cognitive and motor demands are believed to be different. The paper-and-pencil task requires interaction between visuospatial attention and motor action, whereas the computerized task requires comparatively less motor action allowing increased sensitivity to the perceptual senses [18]. Patients with neglect show a reduction in attention towards the ipsilesional space when the perceptual demands are increased [26]. Reduction in attentional capacity with increased processing demands is thought to reflect reduced attention to the ipsilesional space [26]. Less motorically loaded and more perceptually sensitive computerized tasks may result in a bias away from the lesioned hemispace [18].In addition to such differences, previous studies suggest that changes made to the testing modality or procedures may lead to biased results. For example, cues, azimuthal spatial position, modality, and other factors that may alter visuospatial stimulation may influence the testing results [24]. Therefore, maintaining the test methodology may be a vital component in ULN assessment and rehabilitation.

Although the paper-and-pencil testing modality has been the golden standard in spatial neglect assessment for years, the assessment has to be done manually with a ruler and a calculator by an experienced clinical specialist or assessor, allowing intra- and inter-rater variations. In addition, such manual assessment consumes a significant amount of clinical resources such as time and effort of the assessors [12,13]. The semi-computerized e-system with conventional writing tools allows exclusion of the errors derived from the extraneous variables that may result from the altered testing modality and offered automatic real-time assessment for the clinical specialists. As such, the e-system was tested for its feasibility on cognitively healthy participants followed by cognitively impaired patients with percentage deviation as the main assessment variable for its clinical feasibility.

The inter-rater comparison of the e-system showed excellent or nearly perfect correlations, especially for the patients (▶ Table 2). Construct validity of the line bisection test was conducted previous with the star cancellation test (Pearson r=.40, p=.02) and test-retest reliability was reported to be moderate with variable ULN patient responses to line bisection tasks [11]. The star cancellation test is one of the paper-and-pencil tasks for an assessment of a brain lesion and visual neglect [6]. It is commonly conducted with LBTs for modest correlation between two tests to provide a fitter clinical diagnosis. In additional, previously reported results on computerized spatial neglect assessment tools showed similar correlating values between .80 and .88, supporting the inter-rater reliability of the e-system [27]. The test-retest reliability comparisons showed similar reliability (Pearson r=.64, p=.40) in assessments between the e-system and two raters in both healthy participants and patients (▶ Table 1) [11,28]. The test-retest values are significantly lower than that of the inter-rater values (comparison between the e-system to each rater) due to the day-to-day participant variations.

The neglected lines showed different results between the e-system and the raters, especially for the patient group. Reasons for such differences can be inferred to be due to several reasons: e-system error or assessor discrepancy. First, the sensor on the pen tip may have been misaligned with the micro-pattern printed paper during the bisecting procedure. The patients with less dexterity or coordination capacity compared to the cognitively healthy subjects may have held or tilted the pen at an angle greater than the normal healthy participants [29,30]. The built-in sensor of the e-pen was designed to capture the micro-pattern prints at a designated angle. Since it was previously set and tested at an angle for healthy individuals, different hand positioning may have affected the capturing of writing information. Second, the bisecting pressures of the patients were significantly lower than that of the cognitively healthy subjects. Although virtually identifiable by a naked eye, some of the bisecting lines drawn by the patients were made at an intensity normally below the threshold of the pen's sensor. Third, according to the LBT assessment criteria, the bisecting lines are missed or neglected if the bisecting line is not conducted in order from top to bottom. The e-system was designed to catch the neglected lines in accordance with the LBT assessment criteria. The patients,

despite a full explanation of the LBT procedure prior to the test, may have unconsciously bisected the lines without considering the order. Although the assessor monitors the subject during the test, an instantaneous action conducted by the patients may have been missed. Fourth, the e-system was designed to recognize the first bisecting line and count the second line as a missed line. The subjects may have attempted to cross the line once again if the first attempt seemed to be missed. Such action could also be missed by the assessor. Finally, the bisecting lines that uncross the horizontal lines are considered missed, according to the LBT criteria. The e-system counts the under-crossed lines as a missed event. Although most of the subjects draw bisecting lines in a full and complete fashion, some subjects, especially the patients, may draw incomplete lines or lines that barely touch the horizontal lines. The raters may subjectively evaluate such bisecting lines as either missed or unmissed leading to inter-rater and inter-modality variability.

The abovementioned human errors may have occurred. However, the e-system should detect bisected lines detected by a naked eye. Neglect in cognitively healthy subjects is rarely reported. Schenkenberg and colleagues reported that the average neglected rates for the hospital control patients and patients with right unilateral brain damaged were 1.4 and 6.6, respectively, with a maximum of 2 neglected lines per subject [5]. However, cognitively healthy participants showed a similar neglect rate with more than 2 neglected lines in 2 of the participants. The sensitivity of the e-system was needed to identify true bisecting lines detectable by a naked eye. Therefore, the e-system was modified to increase the sensitivity of the drawing information and to exclude system errors. A few minor changes were made to the e-pen. The optical filter was replaced to reduce the effects of the ambient light and reflection light of the IR-LED. Additional infra-red LED was added to double the brightness and identify the patterns regardless of the pen angle. Cognitively healthy participants (HGm) were recruited again to evaluate the modified e-system. The evaluation results showed that the e-system sensitivity was adequately increased to avoid excessively identified neglected lines.

The assessment duration indicates the time it took for either the e-system or the raters to measure and calculate the LBT results completed by the participants (▶Table 1). The results indicated that it took approximately 5 to 6 minutes on the average by the experienced raters to evaluate each test. In addition to assessment, the raters must input the assessed results in the hospital database system.

There are several advantages to the semi-computerized ULN test and assessment system. The e-system instantaneously computed the test results without the intra- or rater variability for quick and reliable accessibility of the test results. The manual assessment could be time consuming for the assessors. Inter- and intra-assessor variations could exist leading to na omission of mild deficits. The degree of ULN, improvement, or deterioration could be unidentified if the mild deficits were not identified [11]. Furthermore, the e-system could provide additional assessment tools and information on UNL. The current e-system can incorporate the deviation direction, pressure, vibration, and pen angle (tilting) information of the user to the LBT program without altering the hardware or assessment duration. Additional ULN assessment tests such as cancellation, bells, or copy and draw tests could be programmed and incorporated into the e-system without altering hardware of the e-system. Additional observations and clinical feedback are needed on the critical assessment requirements of the cognitively impaired patients with ULN to provide better assessment tools for the patients and the clinical specialists.

Limitations

There were some limitations to the study. Based on the observations in working with both the cognitively healthy participants and partially impaired patients with ULN, prominent differences were observed unique to the clinical status of the patients. Although LBT has its advantages, it also has limitations. LBT does not fully diagnose functional deterioration caused by a brain lesion. Performance dissociations between different types of paper-and-pencils tasks have been reported. Such discrepancies indicate that different tasks induce different types of cognitive demands. For example, LBT emphasizes the correct perception of a stimulus whereas a cancellation task emphasizes a normal visual search within a spectrum stimuli [6]. In order to obtain more informative clinical data, the e-system should incorporate a set of test batteries with different assessment information for multi-perspective clinical information.

© Schattauer 2015

H. Jee et al. Feasibility a semi-computerized system for ULN

Another limitation is the cost of the semi-computerized system. The conventional paper-and-pencil format only needed a pencil, printed LBT paper, timer, and ruler. The semi-computerized system requires the e-pen and installable program in additional to a computer. Although many of the conventional assessment tools with analogue devices have been transformed into digital and computer based tools, the cost and the effective of the analysis should always be considered. The current semi-computerized e-system needs to provide more clinically variant information with multiple functions.

In addition, the study was conducted with two different age groups to observe the applicability of the e-system. Although the goal of the study was to observe the feasibility of a novel semi-computerized system, the age of the participants (HG, NG, HGm) should be matched to indicate the clinical significance between the groups. Moreover, although the paper-and-pencil based LBT has been conducted in clinics in Korea for years, a standardized reference guideline has not been reported to compare individual results. Future studies should be designed to offer a reference guideline to compare the degree of ULN. Lastly, the patients recruited in this study do not represent all cognitively impaired patients with ULN. The semi-computerized e-system should be tested on a wider range within a clinical population. A double-blinded study should be conducted on a much larger group of a clinical population which represents all demographic, clinical, and cognitive variables that might affect the applicability and test results. Finally, the study results did not include the directional information with the line bisection scores. The directional information or neglect direction is one of the critical information to assess ULN and indirectly offering information on the location of the lesion.

6. Conclusion

This study was conducted on cognitively healthy individuals to evaluate the general feasibility of a semi-computerized LBT system prior to evaluating the feasibility of the system to the patients diagnosed with unilateral visual neglect. The semi-computerized system (e-system) was both valid and reliable in assessing cognitively healthy and impaired patients. The e-system significantly eliminated the assessment and recording duration by automatically assessing and recording the test results offering efficiency. The e-system can be further modified to capture the unique characteristics of the patients such as the writing pressure and hand tremor information and provide further assistive rehabilitation information.

Clinical Relevance Statement
1. The semi-computerized system simulated the traditional paper-and-pencil LBT assessment without altering the user-friendly environment of the patients
2. The semi-computerized system eliminated the evaluation duration and inter- and intra-rater variability for a more accurate and improved assessment of the patients with unilateral visual neglect.

Conflict of Interest
The authors have no conflict of interest.

Human Subjects Protections
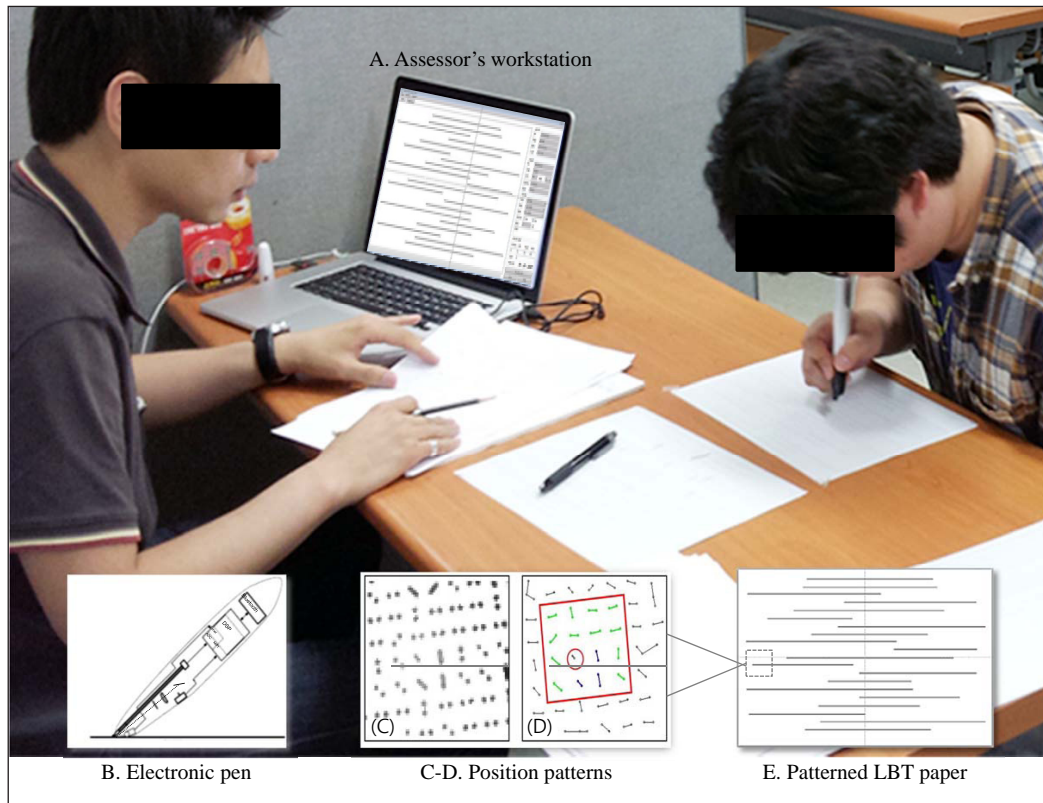The Institutional Review Board of Inha university hospital approved the study protocol.

Fig. 1   E-system assessment configuration for the line-bisection test (LBT)
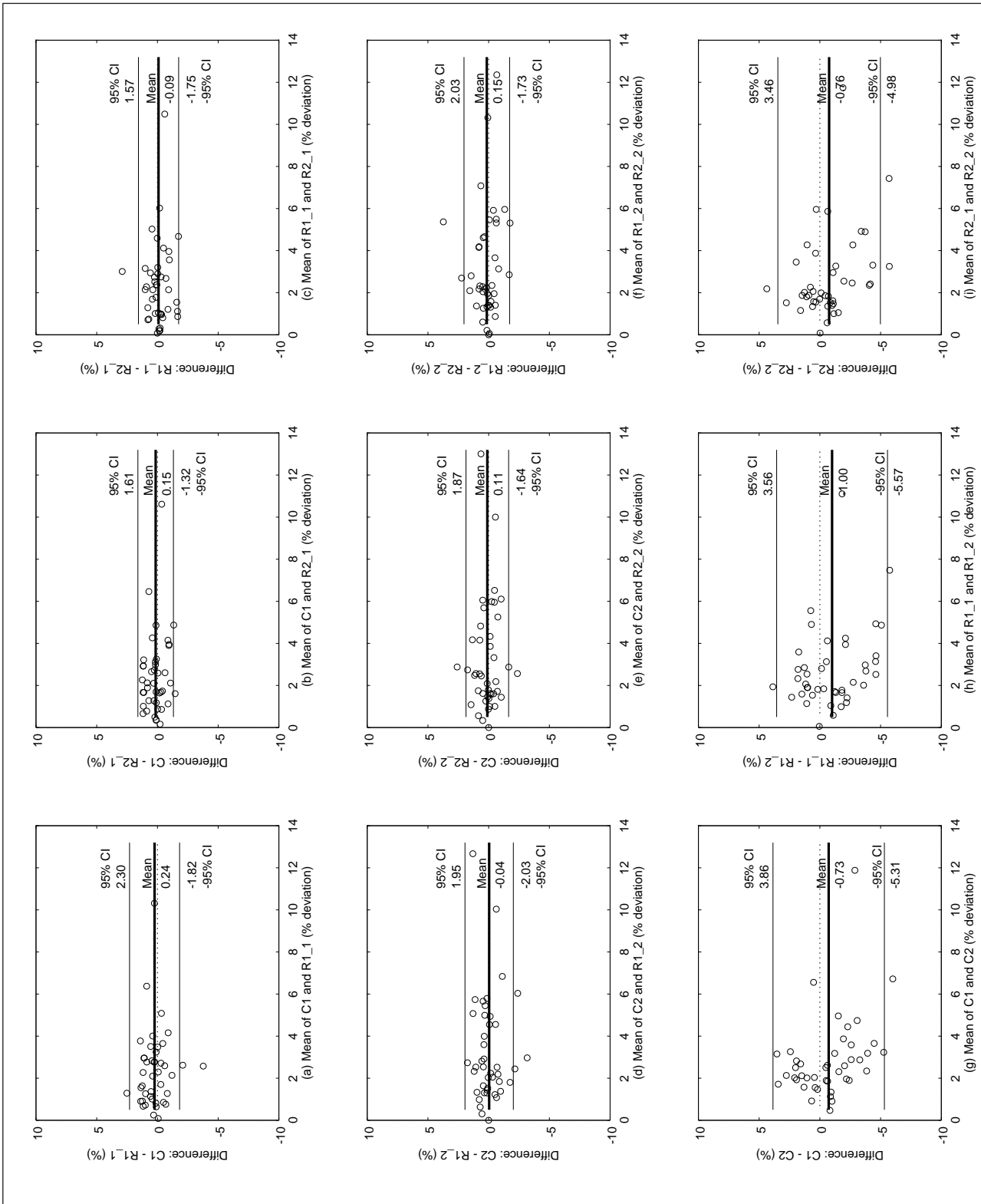
Fig. 2 The Bland-Altman mean difference plots on the percent deviation assessment results obtained during the line bisection test for the healthy participants (HG). Figures (a) through (f) indicate the inter-rater comparisons between the e-system, rater 1, and rater 2. Figures (g) through (l) indicate the intra-rater (test-retest) comparisons between two repeated measures. C1, R1$_1$, and R2$_1$ indicate the percent deviation means from the first set of line bisection test (LBT) assessed by the e-system, rater 1, and rater 2, respectively. C2, R1$_2$, and R2$_2$ indicate the percent deviation means from the second set of LBT assessed by the e-system, rater 1, and rater 2, respectively.
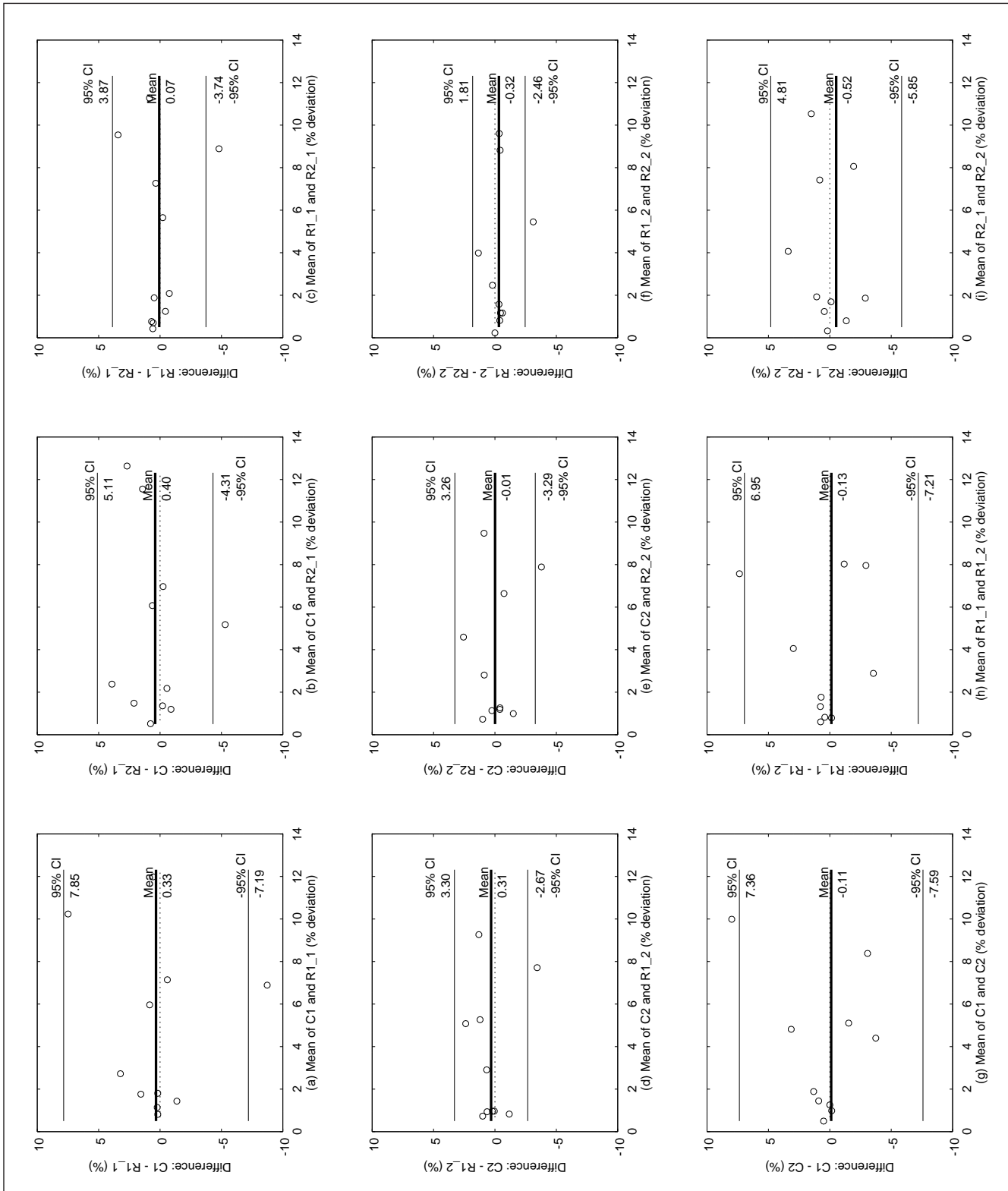
Fig. 3 The Bland-Altman mean difference plots on the percent deviation assessment results obtained during the line bisection test for the patients with unilateral visual neglect (NG). Figures (a) through (f) indicate the inter-rater comparisons between the e-system, rater 1, and rater 2. Figures (g) through (l) indicate the intra-rater (test-retest) comparisons between two repeated measures. C1, R1₁, and R2₁ indicate the percent deviation means from the first set of line bisection test (LBT) assessed by the e-system, rater 1, and rater 2, respectively. C2, R1₂, and R2₂ indicate the percent deviation means from the second set of LBT assessed by the e-system, rater 1, and rater 2, respectively.

Table 1 The intra-rater and inter-rater assessments of the major line bisection test variable, percent deviation, and comparisons of LBT duration, assessment duration, and neglected lines for the cognitively healthy group (HG).

Intra-rater variables (n = 43)	E-system (Mean ± SD)	Rater (Mean ± SD)	p-value	ES	CCC (95% CI)
Test					
LBT duration (sec)	48.29 ± 18.81				
Percent deviation (%) E-system vs. Rater 1	2.51 ± 1.86	2.27 ± 1.95	.15	.12	.84 (.73 to .91)††
Percent deviation (%) E-system vs. Rater 2	2.51 ± 1.86	2.36 ± 1.98	.20	.08	.92 (.86 to .96)††
Percent deviation (%) Rater 1 vs. Rater 2	2.27 ± 1.95	2.36 ± 1.98	.49	.05	.91 (.83 to 0.95)††
Retest					
LBT duration (sec)	64.15 ± 121.74				
Percent deviation (%) E-system vs. Rater 1	3.31 ± 2.58	3.35 ± 2.54	.79	.01	.92 (.86 to .96)††
Percent deviation (%) E-system vs. Rater 2	3.31 ± 2.58	3.20 ± 2.64	.42	.07	.94 (.89 to .97)††
Percent deviation (%) Rater 1 vs. Rater 2	3.35 ± 2.54	3.20 ± 2.64	.30	.06	.93 (.87 to .96)††
Inter-rater variables (n = 43)	**Test (Mean ± SD)**	**Retest (Mean ± SD)**	**p-value**	**ES**	**ICC (95% CI)**
E-system					
Assessment duration (sec)	0	0			
Neglected lines (each)	.93 ± 1.26	.81 ± 1.05			
Percent deviation (%)	2.51 ± 1.86	3.31 ± 2.58	.000	.06	.66 (.37 to .82) ††
Rater 1					
Assessment duration (sec)	325.93 ± 161.70	315.74 ± 130.13			
Neglected lines (each)	0	0			
Percent deviation (%)	2.27 ± 1.95	3.35 ± 2.54	.001	.03	.64 (.34 to .81)††
Rater 2					
Assessment duration (sec)	360.30 ± 157.87	349.36 ± 135.47			
Neglected lines (each)	0	0			
Percent deviation (%)	2.36 ± 1.98	3.20 ± 2.64	.000	.11	.73 (.51 to .86)††

ES: effect size, CI: confidence interval: CCC: concordance correlation coefficient, ICC: intra-class correlation coefficient. The symbols next to the values indicate following meanings: *Significant level of p <0.01, **Significant level, p <0.001.

Table 2 The intra-rater and inter-rater assessments of the major line bisection test variable, percent deviation, and comparisons of LBT duration, assessment duration, and neglected lines for the patient group with unilateral visual neglect (NG).

Intra-rater variables (n = 11)	E-system (Mean ± SD)	Rater (Mean ± SD)	p-value	ES	CCC (95% CI)
Test					
LBT duration (sec)	106.78 ± 28.81				
E-system vs. Rater 1 (%)	5.94 ± 5.28	4.55 ± 4.18	.07	.29	.84 (.59 to .95)**
E-system vs. Rater 2 (%)	5.94 ± 5.28*	4.49 ± 4.24*	.02	.30	.89 (.70 to .96)**
Rater 1 vs. Rater 2 (%)	4.55 ± 4.18	4.49 ± 4.24	.91	.02	.89 (.66 to .97)**
Retest					
LBT duration (sec)	119.19 ± 68.34				
E-system vs. Rater 1 (%)	5.61 ± 5.59*	4.68 ± 5.51	.01	.17	.97 (.90 to .99)**
E-system vs. Rater 2 (%)	5.61 ± 5.59	5.00 ± 5.38	.13	.11	.97 (.89 to .99)**
Rater 1 vs. Rater 2 (%)	4.68 ± 5.51	5.00 ± 5.38	.35	.06	.98 (.92 to .99)**
Inter-rater variables (n = 11)	**Test (Mean ± SD)**	**Retest (Mean ± SD)**	**p-value**	**ES**	**ICC (95% CI)**
E-system					
Assessment duration (sec)	0	0			
Neglected lines (each)	2.27 ± 2.65	2.91 ± 2.55			
Deviation (%)	5.94 ± 5.28	5.61 ± 5.59	.80	.06	.84 (.40 to .96)††
Rater 1					
Assessment duration (sec)	237.91 ± 35.26	231.18 ± 50.86			
Neglected lines (each)	.73 ± 1.68	.73 ± 1.68			
Deviation (%)	4.55 ± 4.18	4.68 ± 5.51	.91	.03	.84 (.41 to .96)††
Rater 2					
Assessment duration (sec)	229.82 ± 36.57	232.27 ± 48.75			
Neglected lines (each)	1.00 ± 2.49	.82 ± 1.94			
Deviation (%)	4.49 ± 4.24	5.00 ± 5.38	.54	.11	.91 (.68 to .98)††

ES: effect size, CI: confidence interval: CCC: concordance correlation coefficient, ICC: intra-class correlation coefficient. The symbols next to the values indicate following meanings: * Significant level of p <0.01, **Significant level, p <0.001.

Table 3   The intra-rater assessments on the major line bisection test variable, percent deviation, and comparisons of assessment duration and neglected lines for the cognitively healthy group with modified e-system (HGm)

Intra-rater variables (HGc = 11)	E-system (Mean ± SD)	Rater (Mean ± SD)	p-value	B-A diff.(95% LOA)	ES	CCC (95% CI)
Percent deviation (%) E-system vs. Rater 1	2.85 ± 3.30	2.93 ± 3.11	.86	.06 (-.96 to .87)	.05	.90 (.67 to .97)**
Assessment duration (sec)	0	0				
Neglected lines (each)	0	0				
Percent deviation (%)E-system vs. Rater 2	2.85 ± 3.30	2.79 ± 3.26	.82	.01 (-.55 to .56)	.11	.96 (.88 to .99)**
Assessment duration (sec)	269.80 ± 107.89	272.20 ± 107.54				
Neglected lines (each)	0	0				
Percent deviation (%)Rater 1 vs. Rater 2	2.93 ± 3.11	2.79 ± 3.26	.73	.06 (-.76 to .89)	.04	.92 (.73 to .98)**
Assessment duration (sec)	283.80 ± 58.16	287.50 ± 65.70				
Neglected lines (each)	0	0				

ES: effect size, CI: confidence interval: CCC: concordance correlation coefficient, ICC: intra-class correlation coefficient, B-A diff.: Bland Altman mean difference or bias, LOA: Limit of agreement (mean difference with 95% CI). The symbols next to the values indicate following meanings: *Significant level of p <0.01, **Significant level, p <0.001.

REFERENCES

1. de Joode E, van Heugten C, Verhey F, van Boxtel M. Efficacy and usability of assistive technology for patients with cognitive deficits: a systematic review. Clin Rehabil 2010; 24(8): 701–714.
2. Cicerone KD, Langenbahn DM, Braden C, Malec JF, Kalmar K, Fraas M, Felicetti T, Laatsch L, Harley JP, Bergquist T, Azulay J, Cantor J, Ashman T. Evidence-based cognitive rehabilitation: updated review of the literature from 2003 through 2008. Arch Phys Med Rehabil 2011; 92(4): 519–530.
3. Jee HM. Evolution of exercise rehabilitation. Journal of Exercise Rehabilitation 2014; 10(3): 343.
4. Unsworth CA. Cognitive and Perceptual Dysfunction. Philadelphia: F.A: Davis Company; 2007.
5. Schenkenberg T, Bradford DC, Ajax ET. Line bisection and unilateral visual neglect in patients with neurologic impairment. Neurology 1980; 30(5): 509–517.
6. Ferber S, Karnath HO. How to assess spatial neglect--line bisection or cancellation tasks? J Clin Exp Neuropsychol 2001; 23(5): 599–607.
7. Molenberghs P, Sale MV. Testing for spatial neglect with line bisection and target cancellation: are both tasks really unrelated? PLoS One 2011 ;6(7): e23017.
8. Ting DS, Pollock A, Dutton GN, Doubal FN, Thompson M, Dhillon B. Visual neglect following stroke: current concepts and future focus. Surv Ophthalmol 2011; 56(2): 114–134.
9. Erez AB, Katz N, Ring H, Soroker N. Assessment of spatial neglect using computerised feature and conjunction visual search tasks. Neuropsychol Rehabil 2009; 19(5): 677–695.
10. Verdon V, Schwartz S, Lovblad KO, Hauert CA, Vuilleumier P. Neuroanatomy of hemispatial neglect and its functional components: a study using voxel-based lesion-symptom mapping. Brain 2010; 133(3): 880–894.
11. Plummer P, Morris ME, Dunai J. Assessment of unilateral neglect. Phys Ther 2003; 83(8): 732–740.
12. Baheux K, Yoshizawa M, Seki K, Handa Y. Virtual reality pencil and paper tests for neglect: a protocol. Cyberpsychol Behav 2006; 9(2): 192–195.
13. Liang Y, Fairhurst MC, Guest RM, Potter JM. A learning model for the automated assessment of hand-drawn images for visuo-spatial neglect rehabilitation. IEEE Trans Neural Syst Rehabil Eng 2010; 18(5): 560–570.
14. Hannaford S, Gower G, Potter JM, Guest RM, Fairhurst MC. Assessing visual inattention: study of inter-rater reliability. Br. J. Therapy Rehabil 2003; 10(2): 72–75.
15. Stone SP, Wilson B, Wroot A, Halligan PW, Lange LS, Marshall JC, Greenwood RJ. The assessment of visuo-spatial neglect after acute stroke. J Neurol Neurosurg Psychiatry 1991; 54(4): 345–350.
16. Gauthier L DF, Joanette Y. The Bells Test: A quantitative and qualitative test for visual neglect. Int J Clin Neuropsychol 1989; 11: 49–54.
17. Deouell LY, Sacher Y, Soroker N. Assessment of spatial attention after brain damage with a dynamic reaction time test. J Int Neuropsychol Soc 2005; 11(6): 697–707.
18. Rolfe MH, Hamm JP, Waldie KE. Differences in paper-and-pencil versus computerized line bisection according to ADHD subtype and hand-use. Brain Cogn 2008; 66(2): 188–195.
19. Ku J, Lee JH, Han K, Kim SI, Kang YJ, Park ES. Validity and reliability of cognitive assessment using virtual environment technology in patients with stroke. Am J Phys Med Rehabil 2009; 88(9): 702–710.
20. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33(1): 159–174.
21. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1(8476): 307–310.
22. McGrath RE, Meyer GJ. When effect sizes disagree: the case of r and d. Psychol Methods 2006; 11(4): 386–401.
23. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
24. Jewell G, McCourt ME. Pseudoneglect: a review and meta-analysis of performance factors in line bisection tasks. Neuropsychologia 2000; 38(1): 93–110.
25. Mendez MF, Cherrier MM, Cymerman JS. Hemispatial neglect on visual search tasks in Alzheimer's disease. Neuropsychiatry Neuropsychol Behav Neurol 1997; 10(3): 203–208.
26. Lavie N, Robertson IH. The role of perceptual load in neglect: rejection of ipsilesional distractors is facilitated with higher central load. J Cogn Neurosci 2001; 13(7): 867–876.
27. Fordell H, Bodin K, Bucht G, Malm J. A virtual reality test battery for assessment and screening of spatial neglect. Acta Neurol Scand 2011; 123(3): 167–174.
28. Kinsella G, Packer S, Ng K, Olver J, Stark R. Continuing issues in the assessment of neglect. Neuropsychol Rehabil 1995; 5: 239–258.

© Schattauer 2015

H. Jee et al. Feasibility a semi-computerized system for ULN

29. Sunderland A, Bowers MP, Sluman SM, Wilcock DJ, Ardron ME. Impaired dexterity of the ipsilateral hand after stroke and the relationship to cognitive deficit. Stroke 1999; 30(5): 949–955.

30. Schaefer SY, Haaland KY, Sainburg RL. Hemispheric specialization and functional impact of ipsilesional deficits in movement coordination and accuracy. Neuropsychologia 2009; 47(13): 2953–2966.